2700/AF

TRANSMITTAL LETTER
(General - Patent Pending)

	Docket No: #18
	02100033AA

6/2/03

In Re Application Of: K. Ozawa

OIPE
MAY 27 2003
PATENT TRADEMARK OFFICE

Serial No. 09/302,397	Filing Date 04/30/1999	Examiner A. Armstrong	Group Art Unit 2654

Title: **SPEECH CODING APPARATUS AND SPEECH DECODING APPARATUS**

<u>TO THE COMMISSIONER OF PATENTS AND TRADEMARKS:</u>

Transmitted herewith is:
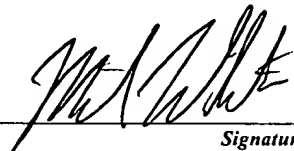
Appellant's Brief Under 37 C.F.R. 1.192 - in triplicate

RECEIVED

MAY 3 0 2003

Technology Center 2600

in the above identified application.

☐ No additional fee is required.

☒ A check in the amount of $320.00 is attached.

☒ The Commissioner is hereby authorized to charge and credit Deposit Account No. 50-2041
as described below. A duplicate copy of this sheet is enclosed.

   ☐ Charge the amount of

   ☒ Credit any overpayment.

   ☒ Charge any additional fee required.

Dated: **May 27, 2003**

Signature

Michael E. Whitham
Reg. No. 32, 635
Whitham, Curtis & Christofferson, P.C.
11491 Sunset Hills Road, Suite 340
Reston, VA 20190
(703) 787-9400

30743
PATENT TRADEMARK OFFICE

CC:

1

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

In re patent application of

Kazunori Ozawa

Serial No.: 09/302,397   Group Art Unit: 2654

Filed: April 30, 1999   Examiner: A. Armstrong

For: SPEECH CODING APPARATUS AND
    SPEECH DECODING APPARATUS

RECEIVED

Box Appeal   **MAY 3 0 2003**
Commissioner for Patents
P.O. Box 1450   Technology Center 2600
Alexandria, VA 22313-1450

APPELLANT'S BRIEF UNDER 37 C.F.R. §1.192

This brief, which is filed herewith in triplicate, is in furtherance of the Notice of Appeal, timely filed in this case on March 24, 2003 from the action of the Examiner finally rejecting claims 1-11 in this application. Attached is a check in amount of $330.00 (37 C.F.R. 1.17(f)) to cover the fee for filing this appeal brief.

This brief contains these items under the following headings, and in the order set forth below (37 C.F.R. §1.192(c)):

I. REAL PARTY IN INTEREST

II. RELATED APPEALS AND INTERFERENCES

III. STATUS OF CLAIMS

IV. STATUS OF AMENDMENTS

V. SUMMARY OF INVENTION

VI. ISSUES

I. REAL PARTY IN INTEREST

The real party in interest in the appeal is:

☐ the party named in the caption of this brief.

☑ the following party: NEC Corporation of Tokyo, Japan.

II. RELATED APPEALS AND INTERFERENCES

With respect to other appeals or interferences that will directly affect, or be directly affected by, or have a bearing on the Board's decision in this appeal:

☑ there are no such appeals or interferences.

☐ these are as follows:

5

III. STATUS OF CLAIMS

The status of the claims in this application are:

A. Total number of claims in the application are: claims 1 to 4 and 6 to 11

B. Status of all the claims:

    1. Claims cancelled: claim 5

    2. Claims withdrawn from consideration but not cancelled: n/a

    3. Claims pending: claims 1 to 4 and 6 to 11

    4. Claims allowed: none

    5. Claims rejected: claims 1 to 4 and 6 to 11

C. Claims on Appeal.

    The claims on appeal are: claims 1 to 4 and 6 to 11

IV. STATUS OF AMENDMENTS

The status of amendments filed subsequent to the final rejection are as follows: An amendment, filed January 23, 2003, as a response to the final rejection has been entered as indicated in the Advisory Action mailed February 20, 2003. That amendment canceled claim 5.

V. Summary of Invention

The invention as defined in the claims on appeal is directed to a speech coding apparatus for coding a speech signal at a low bit rate with high quality. The invention effectively suppresses deterioration in sound quality in terms of background noise while minimizing the calculations required.

A conventional method of coding a speech signal with high efficiency is Code Excited Linear Predictive (CELP) coding. Under the CELP coding scheme, on the transmission side, spectrum parameters representing a spectrum characteristic of a speech signal are extracted from the speech signal for each frame using Linear Predictive Coding (LPC) analysis. Each frame is divided into subframes, and for each subframe, parameters for an adaptive codebook are extracted based on the sound source signal in the past and then the speech signal of the subframe is pitch predicted using the adaptive codebook. This convention coding scheme has the serious disadvantage of requiring a large number of calculations to select an optimum sound source code vector from the sound source codebook.

Various methods have been devised to reduce the calculations required to search a sound source codebook. One such method is the ACELP (Algebraic Code Excited Linear Prediction) method in which the sound source signal is represented by a plurality of pulses and transmitted while the positions of the respective pulses are represented by predetermined numbers of bits. While this approach greatly reduces the calculations required, when background noise is superimposed on speech, the background noise portion of the coded speech greatly reduces the sound quality for bit rates less than around 8kb/s.

The present invention provides a speech coding system which succeeds in reducing the required calculations while at the same time maintaining good sound quality in terms of background noise for low bit rates. Four embodiments of the

speech coding system are shown in Figures 1 to 4, respectively, and a speech decoding system is shown in Figure 5.

With reference to Figure 1 of the drawings, the invention employs a spectrum parameter calculation section which includes a spectrum parameter calculation circuit 200 for extracting spectrum parameter from a speech signal and a spectrum quantizing circuit 210 for quantizing the spectrum parameter. An adaptive codebook section includes adaptive codebook 500 and sound source quantization circuit 350. A mode discrimination circuit 370 discriminates the mode on the basis of the past quantized gain from gain quantization circuit 366. The mode discrimination circuit 370 receives the adaptive codebook gain quantized by the gain quantization circuit 366 one subframe ahead of the current subframe and compares it to a predetermined threshold to perform voiced/unvoiced determination. When a predetermined mode is discriminated, a sound source quantization circuit 350 searches combinations of code vectors stored in a sound source code books 351 or 352, which are used to collectively quantize the amplitudes or polarities of a plurality of pulses, and a plurality of shift amounts used to temporally shift predetermined pulse positions, to select a combination of a code vector and shift amount which minimizes distortion relative to input speech. Codebook 351 is used for voiced sound and codebook 352 is used for unvoiced sound. Appellant uses separate codebooks for voiced and unvoiced speech signals so as to minimize calculations without adversely affecting sound quality in terms of background noise. A multiplexer section 400 outputs a combination of an output from the spectrum parameter calculatioin section, an output from the adaptive codebook section, and an output from the sound source quantization section.

The speech decoding apparatus of the invention is shown in Figure 5 and includes a demultiplexer section 510 for receiving and demultiplexing a spectrum parameter, a delay of an adaptive codebook, a quantized gain, and quantized sound

source information. A mode discrimination section 530 discriminates the mode on the basis of the past quantized gain of the adaptive codebook. A sound decoding section 540 reconstructs a sound source signal by generating non-zero pulses from the quantized sound source information. A speech signal is reproduced or resynthesized by passing the sound source signal through a synthesis filter 560 defined by spectum parameters.

The second embodiment of the coding system according to the invention shown in Figure 2 differs from the first embodiment in the operation of a sound source quantization circuit 355. When voiced/unvoiced discrimination indicates an unvoiced sound, the positions are generated in advance in accordance with a predetermined rule. A random number generating circuit 600 is used to generate a predetermined number of pulse positions which are output to the sound source quantization circuit 355. If the discrimination information indicates a voiced sound, the sound source quantization circuit 355 operates in the same manner as the sound source quantization circuit 350 in Figure 1. If, on the other hand, the information indicates an unvoiced sound, the amplitudes or polarities of pulses are collectively quantized by using a sound source codebook 352 in correspondence with the positions output from the random number generating circuit 600.

In the third embodiment of the coding system according to the invention shown in Figure 3, when voiced/unvoiced discrimination information indicates an unvoiced sound, a sound source quantization circuit 356 calculates the distortions given by equations (21) on page 31 of the specification in correspondence with all the combinations of all the code vectors in a sound source codebook 352 and the shift amounts of pulse positions, selects a plurality of combinations in the order which minimizes distortions, and outputs them to the gain quantization circuit 366. The gain quantization circuit 366 quantizes gains for a plurality of sets of outputs from the sound source quantization circuit 356 by using a codebook 380 and selects a

combination of a shift amount, sound source code vector, and gain code vector which minimizes distortions.

In the fourth embodiment of the coding system according to the invention shown in Figure 4, when voiced/unvoiced discrimination information indicates an unvoiced sound, a sound source quantization circuit 357 collectively quantizes the amplitudes or polarities of pulses for the pulse positions generated by a random number generating circuit 600 by using a sound source codebook 352 and outputs all the code vectors or a plurality of code vector candidates to a gain quantization circuit 367. The gain quantization circuit 367 quantizes gains for the respective candidates output from the sound source quantization circuit 357 by using a gain codebook 380 and outputs a combination of a code vector and gain code vector which minimizes distortion.

A main feature of the present invention as recited in the claims is that a speech coding apparatus comprises a sound source quantization section which has a codebook for representing a sound source signal by a combination of a plurality of non-zero pulses and collectively quantizing amplitudes or polarities of the pulses when an output from the discrimination section indicates a predetermined mode, and searches combinations of code vectors stored in the codebook and a plurality of shift amounts used to shift positions of the pulse so as to output a combination of a code vector and shift amount which minimizes distortion relative to input speech.

Specifically, claim 1 recites, "a sound source quantization section for quantizing a sound source signal of the speech signal by using the spectrum parameter and outputting the sound source signal". This sound source quantization section comprises, "a discrimination section [370] for discriminating a voiced sound mode and an unvoiced sound mode on a basis of a past quantized gain of an adaptive codebook [500] ", and "a sound source quantization section [350] which has a codebook [351, 352] for representing a sound source signal by a combination of a

plurality of non-zero pulses and collectively quantizing amplitudes or polarities of the pulses based on an output from said discrimination circuit section, and searches combinations of code vectors stored in said codebook and a plurality of shift amounts used to shift positions of the pulses so as to output a combination of a code vector and shift amount which minimizes distortion relative to input speech". Similar limitations are recited in claims 2 and 3.

Claim 4 recites, "a sound source quantization section for quantizing a sound source signal of the speech signal by using the spectrum parameter and outputting the sound source signal". The sound source quantization section comprises, "a discrimination section [370] for discriminating a voice sound mode and an unvoiced sound mode on the basis of a past quantized gain of an adaptive codebook [500]", and "a sound source quantization section [350] which has a codebook [351, 352] for representing a sound source signal by a combination of a plurality of non-zero pulses and collectively quantizing amplitudes or polarities of the pulses based on output from said discrimination section, and a gain codebook [380] for quantizing gains, and outputs a combination of a code vector and gain code vector which minimizes distortion relative to input speech by generating positions of the pulses according to a predetermined rule".

Claims 6 and 7 recite the combination of a speech coding and decoding apparatus and, again, recite limitations similar to those pointed out above but, in addition, recite the features of the decoding apparatus of Figure 5. Specifically, claim 6 recites "a demultiplexer section [510] for receiving and demultiplexing a spectrum parameter, a delay of an adaptive codebook, a quantized gain, and quantized sound source information", "a mode discrimination section [530] for discriminating a mode by using a past quantized gain in said adaptive codebook", "a sound source signal reconstructing section [540] for reconstructing a sound source signal by generating non-zero pulses from the quantized sound source information when an output from

said discrimination indicates a predetermined mode", and "a synthesis filter section [560] which is constituted by spectrum parameters and reproduces a speech signal by filtering the sound source signal".

Claim 8 is directed to a speech coding apparatus which comprises, *inter alia,* "mode discrimination means [370] for receiving a past quantized adaptive codebook gain and performing mode discrimination associated with a voiced/unvoiced mode by comparing the gain with a predetermined threshold" and "sound source quantization means [350] for quantizing a sound source signal of the speech signal by using the spectrum parameter and outputting the signal, and searching combinations of code vectors stored in a codebook for collectively quantizing amplitudes or polarities of a plurality of pulses in a predetermined mode and a plurality of shift amounts used to temporally shift a predetermined pulse position so as to select a combination of an index of a code vector and a shift amount which minimizes distortion relative to input speech". Claims 9, 10 and 11 are dependent on claim 8.

VI. ISSUES

The first issue in this appeal is whether the claims 1 to 4 and 6 to 11 contain subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention.

The second issue in this appeal is whether claims 1 to 4 and 6 to 11 are unpatentable over U.S. Patent No. 5,704,003 to Kleijn et al. in view of U.S. Patent No. 5,751,903 to Swaminathan et al. and U.S. Patent No. 5,657,418 to Gerson et al. Put another way, when the claimed invention is considered as a whole, do the patents to Kleijn et al, Swaminathan et al. and Gerson et al., each considered as a whole and viewed without the benefit of impermissible hindsight vision afforded by the claimed invention, suggest the desirability and thus the obviousness of making the combination proposed by the Examiner and is there a reasonable expectation of success in the proposed combination, which is the standard by which obviousness is determined.

VII. GROUPING OF CLAIMS

The claims do not stand or fall together. The claims are grouped as follows: Group 1 includes claims 1-4 and 9-11; Group 2 includes claims 6 and 7.

Claims 1 to 4 are independent claims directed to the several embodiments of the speech encoder shown in Figures 1 to 4. Claim 8 is an independent claim directed to a speech coding apparatus, and claims 9 to 11 are dependent claims dependent on claim 8. More specifically, claim 9 recites that the sound source quantization means uses a position generated according to a predetermined rule as a pulse position when mode discrimination indicates a predetermined mode. Claim 10 recites that, when mode discrimination indicates a predetermined mode, a predetermined number of pulse positions are generated by random number generating means and output to the sound source quantization means. This is shown in the embodiments of Figure 2 and 4. Claim 11 recites *inter alia* that, when mode discrimination indicates a predetermined mode, the sound quantization means selects a plurality of combinations from combinations of all code vectors in the codebook and shift amounts for pulse positions in an order in which a predtermined distortion amount is minimized.

Claims 6 and 7 are independent claims directed to the combination of a coding/decoding apparatus employing the coding system of one of Figures 1 to 4 and the decoding system of Figure 5.

Reasons as to why the grouped claims are separately patentable are included in the Summary of the Invention, *supra*, and the following arguments.

ARGUMENT VIIIA. REJECTIONS UNDER 35 U.S.C. §112, FIRST PARAGRAPH

The Examiner rejected claims 1 to 4 and 6 to 11 under 35 U.S.C. §112, first paragraph, as containing subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and /or use the invention. The Examiner states that "Claims 1 to 11, *as argued by applicant*, include subject matter for processing pulses by using different pulse-shifting schemes depending upon whether a voice [sic] sound mode or n [sic] unvoiced sound mode is discriminated" (emphasis added). The Examiner asserts that the specification does not support a time shift of both voiced and unvoiced sound modes.

The Examiner refers to page 26, lines 14–19, of the specification, stating that the time shift amount $\delta(j)$ is used only in minimizing distortions of unvoiced signals. The Examiner argues that the time shifts are used only in minimizing distortions of unvoiced signals and, therefore, an argument that time shifting is employed for minimizing distortions in both voiced and unvoiced signals is outside the scope of the disclosed invention.

First, it should be noted that the disclosure stands on its own, and any arguments, however interpreted by the Examiner, are not part of the disclosure. Additionally, in rejecting claims the Examiner does not point out which claims and which limitations in the claims are not supported by the specification.

Next, the Examiner misunderstood the invention or relevance of the equations to be the absence a time shifting in the voiced mode. The specification clearly explains on page 24, line 11, that an amplitude codebook or a polarity codebook can be used. The case of a polarity codebook is described wherein codebook 351 is used for voiced sound and codebook 352 is used for unvoiced sound. The point here is that Appellant has used separate codebooks for voiced and unvoiced speech signals so as

to minimize calculations without adversely affecting sound quality in terms of background noise.

Continuing on page 24 of the specification, lines 16 to 21, it is clearly stated that "For a *voiced* sound, the sound source quantization circuit 350 reads out polatity code vectors from the codebook 351, assigns *positions* to the respective code vectors, and selects a combination of a code vector and a *position* which minimizes distortion . . ." (emphasis added). On page 26, lines 5 to 8, it is clearly stated that "For *unvoiced* periods, as indicated in Table 2, pulse positions are set at *predetermined intervals*, and *shift amounts* for shifting the positions of all pulses *are determined in advance* . . . The sound source quantization circuit 350 further receives polarity code vectors from the polarity codebook (sound source codebook) 352, and searches combinations of *all shift amounts* and all code vectors to select a combination of *shift amount* $\delta(j)$ and a code vector gk which minimizes distortion . . ." (emphasis added). The Examiner has misunderstood what the shift amount $\delta(j)$ for unvoiced sound and misconstrued it to mean something different in kind than positions assigned to code vectors for voiced sounds. Shifting positions of pulses, whether for voiced or unvoiced sounds, is time shifting.

The Examiner has failed to point out what limitations in any of the claims are unsupported by the specification as originally filed. The Examiner's rejection appears to be based on his misunderstanding of arguments made during the course of the prosecution and his misinterpretation of the equations presented in the specification.

In view of the above, the invention satisfies the requirements of 35 U.S.C. §112, first paragraph, and the rejection should be reversed.

ARGUMENT VIIIB. REJECTIONS UNDER 35 U.S.C. §112, SECOND PARAGRAPH

There are no rejections under 35 U.S.C. §112, second paragraph.

ARGUMENT VIIIC. REJECTIONS UNDER 35 U.S.C. §102

There are no rejections under 35 U.S.C. §102.

ARGUMENT VIIID. REJECTIONS UNDER 35 U.S.C. §103

Claims 1 to 4 and 6 to 11 were rejected under 35 U.S.C. §103(a) as being unpatentable over U.S. Patent No. 5,704,003 to Kleijn et al. in view of U.S. Patent No. 5,751,903 to Swaminathan et al. and U.S. Patent No. 5,657,418 to Gerson et al.

The primary reference to Kleijn et al. discloses a method of speech coding using Relaxation Code-Excited Linear Predictive (RCELP) techniques which provides a peak-to-average ratio criterion that determines whether or not time shifting of a speech residual signal should be applied within a certain sub-frame. The coders of Kleijn et al. have a characteristic feature of finding a residual signal $r(n)$ from an input speech signal 101, and coding the residual signal $r(n)$ by applying a time shift to the residual signal $r(n)$ with a time warping device and delay line 107. More specifically, a time shift T, which can minimize a differential electric power between an electric power of a signal $r(n-T)$ having a time shift T from the residual signal $r(n)$ and that of a delayed residual signal $r(n-D(n))$ is firstly determined, and then a coding parameter required for the coding is extracted after applying a time shift T to the residual signal $r(n)$.

The Examiner admits that the reference to Kleijn et al. does not teach a voiced/unvoiced mode based on a past quantized gain of an adaptive codebook. However, the Examiner relies on the patent to Gerson et al. for a teaching of this feature.

The patent to Gerson et al. discloses a method of preparation of excitation source gain information for transmission by reducing the sensitivity of the gain bits to errors and speech coder data rates. It is characteristic of the Gerson et al. method to discriminate a coding mode every predetermined section after a speech signal is input and to reduce a bit rate required for a quantization of gain by changing the user of a gain codebook for quantizing gains of an adaptive codebook and a fixed codebook. The change of the gain codebook is dependent on the coding mode.

The Examiner also admits that the primary reference to Kleijn does not teach a multiplexer for the coder or a decoder with demultiplexer and sound source reconstruction. However, the Examiner relies on the patent to Swaminathan for a teaching of this feature.

The patent to Swaminathan et al. (U.S. Patent No. 5,751,903) discloses a multi-mode CELP encoding and decoding method and device which selectively utilizes backward prediction for the short-term predictor parameters and fixed codebook gain of a speech signal. The characteristic of the Swaminathan et al. method is that it discriminates a coding mode every predetermined section after a speech signal is input and changes whether a backward forecast of the fixed codebook gain or not dependent on the coding mode.

Further, the Examiner admits that the primary reference to Kleijn et al. does not teach that the discriminated voiced/unvoiced mode is based on a past quantized gain of an adaptive codebook. Here, again, the Examiner relies on the patent to Gershon et al. for a teaching of this feature.

The Examiner proposes that if the method taught by Kleijn et al. would be implemented with a discrimination of a voiced/unvoiced mode based on a past quantized gain, as taught by Gerson et al., and a multiplexer would be added for the coder or a decoder scheme from Swaminathan et al., the Appellant's invention could be created. In other words, the Examiner has taken three diverse systems and tried to combine them based on Appellant's own disclosure. Appellant respectfully submits that the combination proposed by the Examiner would be inoperable and/or would not result in or make obvious the claimed invention. Furthermore, the combination would not be obvious to one of ordinary skill in the art due to the diverse and complex subject matter being combined and, therefore represents an impermissible hindsight reconstruction of the invention.

On the contrary, the claimed invention has the following characteristics not taught by the combination of Kleijn et al., Gerson et al. and Sawaninathan et al.:

1. After a speech signal is input, a coding mode is discriminated every predetermined section by using a previous adaptive codebook gain to be quantized.

2. In a search of the codebook, which consists of the assemblage of code vectors of non-zero pulses, in a sound source quantization section, the combination of a code vector and a shift amount for shifting a position of the code vector is searched and output. The search is based on the coding mode.

Under these characteristic features of the claimed invention, a speech coding apparatus is equivalent to such a coder as having a large sized codebook. As a result, it becomes possible to increase the degree of freedom of the pulse position and to improve remarkably the sound quality in a low bit rate. There is no disclosure or any suggestion of the second characteristic above to be found in any of the references.

It should be noted that all three references employ different modifications of CELP coding. Specifically, the reference to Kleijn et al. uses so-called Relaxation Code-Excited Linear Predictive (RCELP) techniques, which provides improvement of the coding efficiency by advantageously employing a concept of an adaptive codebook delay trajectory. The patent to Gerson et al. uses multiple coding modes in preparation of excitation source gain information. The patent to Swaminathan et al. is focused on a modified CELP system which uses backward prediction, enabling an input signal to be reconstructed in part by predicting the signal based on the received parameters and the reconstructed signal of the previously decoded frame. Since these references utilize different modifications to achieve different results it would not be obvious to combine them as proposed by the Examiner. For example, if we would use a voiced/unvoiced mode taught by Gerson et al. in the Kleijn et al. system, based on an adaptive codebook delay trajectory set equal to pitch-period trajectory, there would

be unpredictable results. But it should be noted that the reference to Gerson et al. does not disclose a discriminating a voiced/unvoiced mode based on a past quantized gain of an adaptive codebook as Appellant teaches. Gerson et al. clearly states that the selection of coding mode is performed based on periodicity which is different from Appellant's teaching. Therefore, the Gerson et al. reference does not make up to the deficiencies of the primary reference to Kleijn et al.

In order to better understand the context of the present invention it seems useful to explain the general operation of a CELP system. An input speech signal is digitized and filtered to attenuate direct current, hum, or other low frequency contamination before being encoded. Additionally, the input speech signal is buffered into frames to enable linear predictive analysis, which models the amplitudes or polarities of a plurality of pulses shaping effects of the vocal tract. Further, the frames are partitioned into sub-frames for purposes of excitation analysis, which utilizes the two codebooks to model the excitation of each sub-frame of the input speech signal. A vocal tract filter generates speech by filtering a sum of vectors, scaled by gain parameters, selected from the two codebooks. The vectors ultimately used to model the excitation are selected by comparing the differences between the input signal and the speech signal synthesized from the vector sum, taking into account the noise masking properties of the human ear. Specifically, the differences at frequencies at which the error are less important to the human auditory perception are attenuated, while differences at frequencies at which the error are more important are amplified. After testing all possible codebook vectors, the vectors producing the minimal perceptually weighted error energy are selected to model the input speech. A bitstream of data encoding the selected vectors – i.e., the codebook indices and their codebook gains – is multiplexed with the short-term predictor or vocal tract filter parameters, and transmitted to the decoder.

The decoder reconstructs the excitation vectors represented by the codebook

indices, multiplies the vectors by the appropriate gain parameters, and computes the vector sum representing the excitation of the signal, which is then passed through a vocal tract filter to synthesize the speech.

It should be noted that at low bit rates, a relatively small number of bits are available to encode the input speech signal. As a result, conventional CELP systems either have very few bits to encode the parameters or update the parameters very slowly. All these affect a quality of voice in the reconstructed speech.

The claimed invention improves the quality of speech by discriminating the mode in the beginning of encoding. It should noted that the discrimination of the mode by the present invention differs from the approach taught by Gerson et al. and Swaminathan et al. Specifically, Appellant provides a mode discrimination circuit 370 which discriminates the voice on the basis of the past quantized gain of an adaptive codebook (long-term predictor). In contrast, the reference to Swaminathan et al. concentrates on improving the steps of encoding and decoding based on the short-term predictor parameters and the fixed codebook gain of a speech signal in multi-mode regime and recognize the mode by analyzing line spectral frequencies (LSFs) so-called short-term predictor modeling parameters of the vocal tract, and Gerson et al. selects coding mode based on periodicity.

It is clear that the mode recognition used by Swaminathan et al. cannot be applied in Appellant's system for several reasons. First, the reference to Swaminathan et al. uses line spectral frequencies for each discrete segment of the digitized speech signal to represent short term predictor parameters. In contrast, the Appellant's invention operates with amplitudes or polarities of pulses as short term predictor parameters. Second, in order to recognize mode, Swaminathan et al. analyses the following factors of the signal frame: spectral stationarity, pitch stationarity, zero crossing rate, short term level gradient and short term energy. Appellant's system discriminates mode on a basis of a past quantized gain of an adaptive codebook only.

Appellant's invention in order to provide a speech coding system which suppresses a deterioration in sound quality in terms of background noise includes, "a sound source quantization section for quantizing a sound source signal of the speech signal by using the spectrum parameter and outputting the sound source signal". (Claim 1) This sound source quantization section comprises, "a discrimination section [370] for discriminating a voiced sound mode and an unvoiced sound mode on a basis of a past quantized gain of an adaptive codebook [500]", and " a sound source quantization section [350] which has a codebook [351, 352] for representing a sound source signal by a combination of a plurality of non-zero pulses and collectively quantizing amplitudes or polarities of the pulses based on an output from said discrimination circuit section, and searches combinations of code vectors stored in said codebook and a plurality of shift amounts used to shift positions of the pulses so as to output a combination of a code vector and shift amount which minimizes distortion relative to input speech". Similar limitations are recited in each of the other independent claims 2, 3, 4 and 8, and by virtue of their dependency on claim 8, in dependent claims 9, 10 and 11.

In making the rejection, the Examiner acknowledges that Kleijn et al. do not teach discriminating a voiced/unvoiced mode based on a past quantized gain of an adaptive codebook. The Examiner cited Gerson et al. for a teaching that a lag parameter, which reflects the periodicity, is used to select a particular coding mode. From this, the Examiner concludes that it would have been obvious to modify Kleijn et al. to implement discriminating a voiced/unvoiced mode based on past quantized gain. What the Examiner is attempting to do is to ignore the clear teaching of the primary reference and attribute to the secondary reference a teaching which is clearly not warranted by the reference. The combination which was proposed is based solely on hindsight not permitted by Section 103 of the Patent Statute.

The Examiner states that Kleijn et al. teaches sound source quantization by searching a codebook for code vectors and delays so as to output a combination of code vector and shift amount that minimizes distortion. This in not an accurate characterization of the reference. A characteristic feature of Kleijn et al. is that residual signals are coded by a time shift. That is, as described in column 6, after line 14, the best value for time shift T which can minimize an error output between a signal r(n-T) obtained by shifting the residual signal r(n) by T and a delayed residual signal r(n-D9n)) is required, whereby the parameter required in coding is selected.

The Examiner also acknowledges that Kleijn et al. do not teach a multiplexer for the coder or decoder scheme, but cites Swaminathan et al, saying that "it would have been obvious to implement a multiplexer and a decoding scheme with the system of Kleijn et al. for purpose of providing high quality speech coding and decoding as suggested by Swaminathan et al."

Independent claims 6 and 7 are directed to the combination of a speech coding/decoding apparatus. The coding part of the apparatus is illustrated in the embodiments shown in Figures 1 to 4, while the decoding part of the apparatus is illustrated in Figure 5. This combination is separately patentable from the speech encoding apparatus that is a subcombination of the claimed combination. More specifically, claim 6 recites " a sound source quantization section [350] for quantizing a sound source signal of the speech signal by using the spectrum parameter and outputting the sound source signal", "a discrimination section [370] for discriminating a voice sound mode and an unvoiced sound mode on the basis of a past quantized gain of a adaptive codebook", and "a codebook [350, 351] for representing a sound source signal by a combination of a plurality of non-zero pulses and collectively quantizing amplitudes or polarities of the pulses when an output from said discrimination section indicates a predetermined mode". Claim 6 further recites that "said sound source quantization section [350] . . . [searches] combinations of

code vectors stored in said codebook and a plurality of shift amounts used to shift positions of the pulses so as to output a combination of a code vector and shift amount which minimizes distortion relative to input speech". The speech encoding apparatus also includes "a multiplexer section [400] for outputting a combination of an output from said spectrum parameter calculation section, an output from said adaptive codeboook section, and an output from said sound source quantization section". The speech decoding apparatus includes "a demultiplexer section [510] for receiving and demultiplexing a spectrum parameter, a delay of an adaptive codebook, a quantized gain, and quantized sound source information". ". . . [A] mode discrimination section [530] . . . [discriminates] a mode by using a past quantized gain in said adaptive codebook" and "a sound source signal reconstructing section [540] . . . [reconstructs] a sound source signal by generating non-zero pulses from the quantized sound source information when an output from said discrimination indicates a predetermined mode". Finally, "a synthesis filter section [560] which is constituted by spectrum parameters . . . reproduces a speech signal by filtering the sound source signal." Nothing like this combination is even remotely hinted at by Swaminathan et al., either singly or in combination with Kleijn et al. and Gerson et al. Similar limitations are recited in claim 7.

Appellant respectfully submits that the combination of the prior art proposed by the Examiner would not be functional for the reason that the combined elements of different systems are intended to function for different CELP coding schemes and are not going to be functional in the combination proposed by the Examiner. Even more, the combination of a Relaxation Code-Excited Linear Predictive (RCELP) coding, using a long-term predictor set to equal a pitch-period from Kleijn et al., with a coding mode indicator for the short-term predictor parameters from Gerson et al. and multiplexor form Swaminathan et al. is not the apparatus claimed by the Appellant.

The Examiner has failed to make out a *prima facie* case of obviousness. More specifically, the Examiner has not applied the basic considerations which apply to obviousness rejections as set out in MPEP 2141. Specifically, "When applying 35 U.S.C. 103, the following tenets of patent law must be adhered to:

"(A) The claimed invention must be considered as a whole;

"(B) The references must be considered as a whole and must suggest the desirability and thus the obviousness of making the combination;

"(C) The references must be viewed without the benefit of impermissible hindsight vision afforded by the claimed invention; and

"(D) Reasonable expectation of success is the standard with which obviousness is determined."

The Examiner's rejection is based on (a) a consideration of the parts of the claimed invention, (b) parts of the references without any suggestion for the desirability of making the combination, (c) impermissible hindsight vision afforded by the claimed invention, and (d) without any reasonable expectation of success. Therefore, it is respectfully submitted that rejection under 35 U.S.C. §103 should be reversed.

ARGUMENT VIIIE. REJECTIONS OTHER THAN 35 U.S.C. §§102, 103 AND 112

There are no rejections other than the rejections under 35 U.S.C. §§103 and 112, first paragraph.

IX. APPENDIX OF CLAIMS INVOLVED IN THE APPEAL (37 C.F.R. §1.192(C)(9))

The text of the claims involved in the appeal are:

1   1. A speech coding apparatus including at least:

2   a spectrum parameter calculation section for receiving a speech signal,

3   obtaining a spectrum parameter, and quantizing the spectrum parameter,

4   an adaptive codebook section for obtaining a delay and a gain from a past

5   quantized sound source signal by using an adaptive codebook, and obtaining a residue

6   by predicting a speech signal, and

7   a sound source quantization section for quantizing a sound source signal of the

8   speech signal by using the spectrum parameter and outputting the sound source signal,

9   comprising:

10   a discrimination section for discriminating a voiced sound mode and an

11   unvoiced sound mode on a basis of a past quantized gain of an adaptive codebook;

12   a sound source quantization section which has a codebook for representing a

13   sound source signal by a combination of a plurality of non-zero pulses and

14   collectively quantizing amplitudes or polarities of the pulses based on an output from

15   said discrimination section, and searches combinations of code vectors stored in said

16   codebook and a plurality of shift amounts used to shift positions of the pulses so as to

17   output a combination of a code vector and shift amount which minimizes distortion

18   relative to input speech; and

19   a multiplexer section for outputting a combination of an output from said

20   spectrum parameter calculation section, an output from said adaptive codebook

21   section, and an output from said sound source quantization section.

1   2. A speech coding apparatus including at least:

2   a spectrum parameter calculation section for receiving a speech signal,

3   obtaining a spectrum parameter,

4   an adaptive codebook section for obtaining a delay and a gain from a past

5   quantized sound source signal by using an adaptive codebook, and obtaining a residue

6   by predicting a speech signal, and

7   a sound source quantization section for quantizing a sound source signal of the

8   speech signal by using the spectrum parameter and outputting the sound source signal,

9   comprising:

10   a discrimination section for discriminating a voice soundmode and an

11   unvoiced sound mode on a basis of a past quantized gain of an adaptive codebook;

12   a sound source quantization section which has a codebook for representing a

13   sound source signal by a combination of a plurality of non-zero pulses and

14   collectively quantizing amplitudes or polarities of the pulses based on an output from

15   said discrimination section, and outputs a code vector that minimizes distortion

16   relative to input speech by generating positions of the pulses according to a

17   predetermined rule; and

18   a multiplexer section for outputting a combination of an output from said

19   spectrum parameter calculation section, an output from said adaptive codebook

20   section, and an output from said sound source quantization section.

1   3. A speech coding apparatus including at least:

2   a spectrum parameter calculation section for receiving a speech signal,

3   obtaining a spectrum parameter, and quantizing the spectrum parameter,

4   an adaptive codebook section for obtaining a delay an a gain from a pst

5   quanized sound source signal by using an adaptive codebook, and obtaining a reidue

6   by predicting a speech signal, and

1   a sound source quantization section for quantizing a sound source signal of the

2   speech signal by using the spectrum parameter and outputting the sound source signal,

3   comprising:

4   a discrimination section for discriminating a voice sound mode and an

5   unvoiced sound mode on the basis of a past quantized gain of an adaptive codebook;

6   a sound source quantization section which has a codebook for representing a

7   sound source signal by a combination of a plurality of non-zero pulses and

8   collectively quantizing amplitudes or polarities of the pulses based an output from

9   said discrimination section, and a gain codebook for quantizing gains, and searches

10   combinations of code vectors stored in said codebook, a plurality of shift amounts

11   used to shift positions of the pulses, and gain code vectors stored in said gain

12   codebook so as to output a combination of a code vector, shift amount, and gain code

13   vector which minimizes distortion relative to input speech; and

14   a multiplexer section for outputting a combination of an output from said

15   spectrum parameter calculation section, an output from said adaptive codebook

16   section, and an output from said sound source quantization section.

1   4. A speech coding apparatus including at least:

2   a spectrum parameter calculation section for receiving a speech signal,

3   obtaining a spectrum parameter, and quantizing the spectrum parameter,

4   an adaptive codebook section for obtaining a delay an a gain from a past

5   quantized sound source signal by using an adaptive codebook, and obtaining a residue

6   by predicting a speech signal, and

7   a sound source quantization section for quantizing a sound source signal of the

8   speech signal by using the spectrum parameter and outputting the sound source signal,

9   comprising:

10   a discrimination section for discriminating a voice sound mode and an

11   unvoiced sound mode on the basis of a past quantized gain of an adaptive codebook;

12   a sound source quantization section which has a codebook for representing a

13   sound source signal by a combination of a plurality of non-zero pulses and

14   collectively quantizing amplitudes or polarities of the pulses based on an output from

15   said discrimination section indicates a predetermined mode, and a gain codebook for

16   quantizing gains, and outputs a combination of a code vector and gain code vector

17   which minimizes distortion relative to input speech by generating positions of the

18   pulses according to a predetermined rule; and

19   a multiplexer section for outputting a combination of an output from said

20   spectrum parameter calculation section, an output from said adaptive codebook

21   section, and an output from said sound source quantization section.

1   6. A speech coding/decoding apparatus comprising:

2   a speech coding apparatus including:

3   a spectrum parameter calculation section for receiving a speech signal,

4   obtaining a spectrum parameter, and quantizing the spectrum parameter,

5   an adaptive codebook section for obtaining a delay and a gain from a past

6   quantized sound source signal by using an adaptive codebook, and obtaining a residue

7   by predicting a speech signal,

8   a sound source quantization section for quantizing a sound source signal of the

9   speech signal by using the spectrum parameter and outputting the sound source signal,

10   a discrimination section for discriminating a voice sound mode and an

11   unvoiced sound mode on the basis of a past quantized gain of a adaptive codebook,

12   and

13   a codebook for representing a sound source signal by a combination of a

14   plurality of non-zero pulses and collectively quantizing amplitudes or polarities of the

15   pulses when an output from said discrimination section indicates a predetermined

16   mode,

17   said sound source quantization section searching combinations of code vectors

18   stored in said codebook and a plurality of shift amounts used to shift positions of the

19   pulses so as to output a combination of a code vector and shift amount which

20   minimizes distortion relative to input speech, and further including

21   a multiplexer section for outputting a combination of an output from said

22   spectrum parameter calculation section, an output from said adaptive codeboook

23   section, and an output from said sound source quantization section; and

24   a speech decoding apparatus including at least:

25   a demultiplexer section for receiving and demultiplexing a spectrum

26   parameter, a delay of an adaptive codebook, a quantized gain, and quantized sound

27   source information,

28   a mode discrimination section for discriminating a mode by using a past

29   quantized gain in said adaptive codebook,

30   a sound source signal reconstructing section for reconstructing a sound source

31   signal by generating non-zero pulses from the quantized sound source information

32   when an output from said discrimination indicates a predetermined mode, and

33   a synthesis filter section which is constituted by spectrum parameters and

34   reproduces a speech signal by filtering the sound source signal.

1   7. A speech coding/decoding apparatus comprising:

2   a speech coding apparatus including:

3   a spectrum parameter calculation section for receiving a speech signal,

4   obtaining a spectrum parameter, and quantizing the spectrum parameter,

5   an adaptive codebook section for obtaining a delay and a gain from a past

6   quantized sound source signal by using an adaptive codebook, and obtaining a residue

7   by predicting a speech signal,

8   a sound source quantization section for quantizing a sound source signal of the

9   speech signal by using the spectrum parameter and outputting the sound source signal,

10   a discrimination section for discriminating a voice sound mode and an

11   unvoiced sound mode on the basis of a past quantized gain of an adaptive codebook,

12   and

13   a codebook for representing a sound source signal by a combination of a

14   plurality of non-zero pulses and collectively quantizing amplitudes or polarities of the

15   pulses based on an output from said discrimination section,

16   said sound source quantization section outputting a combination of a code

17   vector and shift amount which minimizes distortion relative to input speech by

18   generating positions of the pulses according to a predetermined rule, and further

19   including

20   a multiplexer section for outputting a combination of an output from said

21   spectrum parameter calculation section, an output from said adaptive codebook

22   section, and an output from said sound source quantization section; and

23   a speech decoding apparatus including at least:

24   a demultiplexer section for receiving and demultiplexing a spectrum

25   parameter, a delay of an adaptive codebook, a quantized gain, and quantized sound

26   source information,

27   a mode discrimination section for discriminating a mode by using a past

28   quantized gain in said adaptive codebook,

29   a sound source signal reconstructing section for reconstructing a sound source

30   signal by generating positions of pulses according to a predetermined rule and

31   generating amplitudes or polarities for the pulses from a code vector when an output

32   from said discrimination section indicates a predetermined mode, and

33   a synthesis filter section which includes spectrum parameters and reproduces a

34   speech signal by filtering the sound source signal.

1   8. A speech coding apparatus comprising:

2   a spectrum parameter calculation section for receiving a speech signal,

3   obtaining a spectrum parameter, and quantizing the spectrum parameter;

4   means for obtaining a delay and a gain from a past quantized sound source

5   signal by using an adaptive codebook, and obtaining a residue by predicting a speech

6   signal; and

7   mode discrimination means for receiving a past quantized adaptive codebook

8   gain and performing mode discrimination associated with a voiced/unvoiced mode by

9   comparing the gain with a predetermined threshold, and

10   further comprising:

11   sound source quantization means for quantizing a sound source signal of the

12   speech signal by using the spectrum parameter and outputting the signal, and

13   searching combinations of code vectors stored in a codebook for collectively

14   quantizing amplitudes or polarities of a plurality of pulses in a predetermined mode

15   and a plurality of shift amounts used to temporally shift a predetermined pulse

16   position so as to select a combination of an index of a code vector and a shift amount

17   which minimizes distortion relative to input speech;

18   gain quantization means for quantizing a gain by using a gain codebook; and

19   multiplex means for outputting a combination of outputs from said spectrum

20   parameter calculation means, said adaptive codebook means, said sound source

21   quantization means, and said gain quantization means.

1   9. An apparatus according to claim 8, wherein said sound source quantization means

2   uses a position generated according to a predetermined rule as a pulse position when

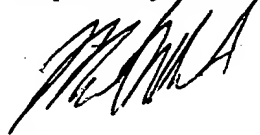3   mode discrimination indicates a predetermined mode.

1 10. An apparatus according to claim 9, wherein when mode discrimination indicates a
2 predetermined mode, a predetermined number of pulse positions are generated by
3 random number generating means and output to said sound source quantization
4 means.

1 11. An apparatus according to claim 8, wherein when mode discrimination indicates a
2 predetermined mode, said sound source quantization means selects a plurality of
3 combinations from combinations of all code vectors in said codebook and shift
4 amounts for pulse positions in an order in which a predetermined distortion amount is
5 minimized, and outputs the combinations to said gain quantization means, and
6   said gain quantization means quantized a plurality of sets of outputs from said
7 sound source quantization means by using said gain codebook, and selects a
8 combination of a shift amount, sound source code vector, and gain code vector which
9 minimizes the predetermined distortion amount.

X. OTHER MATERIALS THAT APPELLANT CONSIDERS NECESSARY OR DESIRABLE

There are no other materials that Appellant considers either necessary or desirable for purposes of this appeal.

Respectfully submitted,

Michael E. Whitham
Reg. No. 32,635

Whitham, Curtis & Christofferson, P.C.
11491 Sunset Hills Road, Suite 340
Reston, VA 20190
Tel. (703) 391-2510
Fax. (703) 391-9035
Customer No. 30743

30743
PATENT TRADEMARK OFFICE